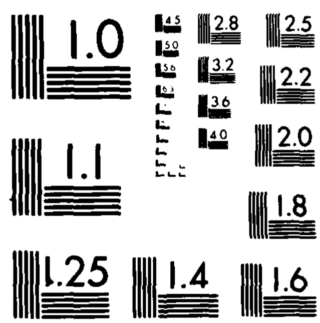
AD-A111 690 CALIFORNIA UNIV BERKELEY OPERATIONS RESEARCH CENTER F/G 12/1
A BAYESIAN INTERPRETATION OF DATA TRIMMING TO REMOVE EXCESS CLA--ETC(U)
DEC 81 W S JEWELL, H BUEHLMANN, A GISLER AFOSR-81-0122
UNCLASSIFIED ORC-81-24 NL

END
DATE
FILMED
4-82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ORC 81-24
DECEMBER 1981

TWO PAPERS ON BAYESIAN DATA TRIMMING

by

WILLIAM S. JEWELL

OPERATIONS
RESEARCH
CENTER

UNIVERSITY OF CALIFORNIA · BERKELEY

TWO PAPERS ON BAYESIAN DATA TRIMMING†

A BAYESIAN INTERPRETATION OF DATA TRIMMING TO REMOVE EXCESS CLAIMS

by

William S. Jewell

EXCESS CLAIMS AND DATA TRIMMING IN THE CONTEXT OF CREDIBILITY RATING PROCEDURES

by

Hans Bühlmann, Alois Gisler and William S. Jewell

DECEMBER 1981 ORC 81-24

†Reproduction was supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-81-0122. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ORC 81-24	2. GOVT ACCESSION NO. AD-A111690	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TWO PAPERS ON BAYESIAN DATA TRIMMING		5. TYPE OF REPORT & PERIOD COVERED Research Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William S. Jewell		8. CONTRACT OR GRANT NUMBER(s) AFOSR-81-0122
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operations Research Center University of California Berkeley, California 94720		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS United States Air Force Air Force Office of Scientific Research Bolling Air Force Base, D.C. 20332		12. REPORT DATE December 1981
		13. NUMBER OF PAGES 40
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Data Trimming
Bayesian Forecasting
Credibility Theory
Outlier Models

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(SEE ABSTRACT)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Foreword

The joint research in these two related papers was supported by the Mathematics Research Institute, Federal Institute of Technology, Zürich, where the University of California author was a visiting scholar during 1980-1981.

They are reproduced in this format solely to facilitate wider distribution before publication. "Excess Claims and Data Trimming in the Context of Credibility Rating Procedures" has been accepted for publication in the Bulletin of the Association of Swiss Actuaries.

A BAYESIAN INTERPRETATION OF DATA TRIMMING TO REMOVE EXCESS CLAIMS

William S. Jewell
University of California, Berkeley
& ETH-Zürich

Abstract

The effect of excess or catastrophic claims is well recognized in insurance. For example, in experience rating it is customary to truncate the data to minimize the effect of such outliers; Gisler has recently proposed a credibility formula using such data trimming. This paper develops a model of the excess claims process and finds the exact Bayesian forecast. The resulting forecast form is approximately a data trim, thus justifying the simpler, heuristic approach.

Zürich, March, 1981

1. Introduction

The effect of excess or catastrophic claims is well recognized in insurance. Typically, one wishes not only to analyze them in detail to determine and, if possible, correct their causes, but also to modify the data so as to minimize their effect upon normal operating procedures of the firm.

For example, in experience rating, data $\underline{x} = (x_1, x_2, \ldots x_n)$ collected from a policyholder's experience in years $1,2,\ldots n$ is used to modify his premium for year $n+1$. If $\tilde{y} = \tilde{x}_{n+1}$ is the random variable denoting next year's total paid claims, the fair premium will be just the regression of $\tilde{y}$ on the data $\underline{x}$, or $\mathcal{E}(\tilde{y}|\underline{x})$. In credibility theory, it is assumed that this forecast is linear in the data, giving the well-known formula:

$$\mathcal{E}(\tilde{y}|\underline{x}) \approx f(\underline{x}) = (1-Z_n)m + Z_n\left(\frac{1}{n}\sum_{t=1}^{n} x_t\right) . \tag{1.1}$$

Here m is the "manual" (fair, no-data) premium, and $Z_n = n/(n+N)$ is the credibility factor with time constant N determined empirically or from a Bayesian model (see, e.g., Norberg(1979) for further details).

The effect of an excess claim upon experience rating is obvious from (1.1). What one would like to do is to detect and remove this claim from the data, and spread all or a portion of the excess amount over all the policyholders, perhaps by charging it against a special reserve. However, in many situations it is not possible or economical to use qualitative information about the claim to decide if it is of ordinary or excess type, and one must use a numerical procedure to "cleanse" the data before using (1.1). Based upon heuristic methods used in industry, A. Gisler (1980) proposed to replace (1.1) by:

$$f(\underline{x}) = a + b\left(\frac{1}{n}\sum_{t=1}^{n} \min(x_t, M)\right) \quad , \tag{1.2}$$

where the parameters (a,b,M) are adjusted so as to minimize the mean-squared error in the forecast; the result could be called a data-trimmed credibility formula.

2. A Bayesian Model for Outliers

We now develop a model which describes how excess claims arise, and then find the exact Bayesian prediction formula. By comparing this with Gisler's form (1.2), we will be able to provide additional motivation for the trimming procedure.

First of all, we assume that an ordinary claim random variable, $\tilde{x}_o$, has a known density, $p_o(x_o|\theta)$, depending upon an unknown parameter θ which characterizes the different policyholders and their exposure characteristics. The first two moments of this random variable are:

$$m_o(\theta) = \mathcal{E}(\tilde{x}_o|\theta) \quad ; \quad v_o(\theta) = \mathcal{V}(\tilde{x}_o|\theta) \ . \tag{2.1}$$

In the usual experience rating model, we are given several independent observations of the $\tilde{x}_o$ type from a policy with fixed, but unknown, θ, and we wish to estimate the mean of the next observation from the same policy. This is equivalent to estimating $m_o(\theta)$, given a prior density on θ, and the data $\underline{x}$. If it is known that all data is of ordinary type, then in many cases the credibility forecast (1.1) is exact, or a good approximation.

Now, however, suppose that it is occasionally possible that we observe instead an excess claim random variable, $\tilde{x}_e$, with density $p_e(x_e)$ not depending upon θ (although this can be easily generalized, if desired). This excess claim is considered to be the result of some extraordinary cause, so that the density p_e will have large mean and variance compared with every density p_o. We also assume that there is no qualitative way in which one can identify an excess claim as such; thus, the densities should have overlapping ranges, otherwise, there would be no difficulty in separating excess claims based upon their magnitude.

We continue to let $\underline{x} = (x_1, x_2, \ldots x_n)$ represent the observational data, assumed independent, given θ. But the observation random variable, $\tilde{x}_t$, $(t=1,2,\ldots n)$, is now sometimes an ordinary, sometimes an excess random variable, and we assume that there is a known contamination probability, π, that independently selects if an ordinary claim is replaced by an excess claim. In other words, we assume that the individual observations follow the mixed density:

$$p(x_t|\theta) = (1-\pi)p_o(x_t|\theta) + \pi p_e(x_t) \quad , \tag{2.2}$$

so that the likelihood of $\underline{x}$, *

$$p(\underline{x}|\theta) = \prod_{t=1}^{n} p(x_t|\theta) \quad , \tag{2.3}$$

consists of 2^n terms. Since π is small, however, only the first few terms will generally be significant (e.g., there are usually only no, one, or a few excess claims in any small sample).

As in other experience rating models, we assume that we are given a prior density, $p(\theta)$, on the unknown parameter, so that Bayes' law then gives a posterior-to-data density for the unknown parameter of:

$$p(\theta|\underline{x}) = \frac{p(\underline{x}|\theta)\ p(\theta)}{p(\underline{x})} \quad , \tag{2.4}$$

*From this point on, we are using the usual Bayesian trick of using p(.) for several different densities, and letting the variables "speak for themselves".

where $p(\underline{x})$ is the integral of the numerator over θ. Now, however, we must remember that we are not interested in predicting just the next observation, but rather in predicting the next observation, <u>given that it is of ordinary type</u>; this random variable, call it $\tilde{y}_o$, has a density $p_o(y_o|\theta)$ if θ were known. It follows then that, given the data $\underline{x}$, we can form the Bayesian <u>predictive density</u> of $\tilde{y}_o$ from (2.4) as follows:

$$(2.5) \qquad p(y_o|\underline{x}) = \int p_o(y_o|\theta)\, p(\theta|\underline{x})\, d\theta \quad .$$

The exact Bayesian <u>mean predictor</u> of y_o is then just the first moment:

$$(2.6) \qquad \mathcal{E}(\tilde{y}_o|\underline{x}) = f(\underline{x}) = \int m_o(\theta)\, p(\theta|\underline{x})\, d\theta \quad .$$

To better understand how this formula depends upon the data, we need to develop further the likelihood (2.3).

3. Single Observation Case

First, suppose that only n=1 observation has been made. Then (2.3) has only two terms, and the exact forecast is:

$$(3.1) \quad f(x_1) = \frac{1}{p(x_1)}\left[(1-\pi)\int m_o(\theta)\, p_o(x_1|\theta)\, p(\theta)\, d\theta + \pi\, p_e(x_1)\int m_o(\theta)\, p(\theta)\, d\theta\right] ,$$

where

$$(3.2)\ p(x_1) = (1-\pi)p_o(x_1) + \pi p_e(x_1) \ ; \ p_o(x_1) = \int p_o(x_1|\theta)p(\theta)d\theta \ .$$

The second integral in (3.1) is just the a priori expected value of $\tilde{y}_o$ (the manual premium):

$$m_o = \mathcal{E} m_o(\tilde{\theta}) = \mathcal{E}\tilde{y}_o \, , \tag{3.3}$$

which would be the "forecast" if no data were available.

The first integral in (3.1) is more interesting, as it is related to the Bayesian prediction in which <u>it is known that the observation is of ordinary type</u>. In contrast to $f(x_1)$, which is the prediction from an arbitrary observation following (2.2), we can <u>define</u> $f_o(x_1)$ as the <u>ordinary observation prediction</u>, gotten from (2.6) by setting $\pi=0$:

$$\mathcal{E}(\tilde{y}_o | x_1 \text{ ordinary}) = f_o(x_1) = \int m_o(\theta) \frac{p_o(x_1|\theta)\, p(\theta)}{p_o(x_1)} \, d\theta \, . \tag{3.4}$$

This could, of course, follow the linear credibility law (1.1).

Finally, we rewrite the exact forecast as:

$$f(x_1) = \frac{1}{p(x_1)} \left[(1-\pi) p_o(x_1) f_o(x_1) + \pi p_e(x_1) m_o \right] , \tag{3.5}$$

which can be rewritten in two revealing forms: the first,

$$f(x_1) = \frac{f_o(x_1) + \phi(x_1) m_o}{1 + \phi(x_1)} \, , \tag{3.6}$$

with

$$\phi(x_1) = \left(\frac{\pi}{1-\pi} \right) \frac{p_e(x_1)}{p_o(x_1)} \tag{3.7}$$

as an "odds-likelihood-ratio"; the second in a credibility format:

$$f(x_1) = \left[1 - Z(x_1) \right] m_o + Z(x_1) f_o(x_1) \, , \tag{3.8}$$

with a new data-dependent credibility factor:

$$(3.9) \qquad Z(x_1) = \left[1+\phi(x_1)\right]^{-1} = \frac{(1-\pi)p_o(x_1)}{(1-\pi)p_o(x_1)+\pi p_e(x_1)} \quad .$$

$Z(x_1)$ is essentially the a posteriori probability that the observation x_1 is ordinary.

In the usual situation, the averaged ordinary density $p_o(x_1)$ and the excess density $p_e(x_1)$ might appear as in Figure 1, giving then the weighting functions $\phi(x_1)$ or $Z(x_1)$ shown in Figure 2.

4. Comparison with Trimming in the Credibility Case

As discussed in the first Section, it is often the case that $f_o(\underline{x})$ is linear in the data $\underline{x}$, i.e. it follows (1.1), with m replaced by m_o, and the ordinary (non-data-dependent) credibility factors Z_n replaced by:

$$(4.1) \qquad Z_{on} = \frac{n}{n+N_o} \quad ; \qquad N_o = \frac{\mathcal{E}v_o(\tilde{\theta})}{\mathcal{V}m_o(\tilde{\theta})} \quad .$$

Thus, in the one-dimensional case, $f_o(x_1)$ in (3.8) would be replaced by:

$$(4.2) \qquad f_o(x_1) = m_o + Z_{o1}(x_1-m_o) \quad .$$

This means that the exact Bayesian forecast would have the interesting shape shown in Figure 3.

This shape shows us that, if x_1 is small, we believe it is of ordinary type, and we should experience rate according to the linear law (4.2). But as x_1 increases beyond m_o into the region where the odds-likelihood-ratio becomes significant, we begin to hedge our bets on the fact that we have an ordinary observation, and to reduce the dependence of the forecast on x_1. Finally, for x_1

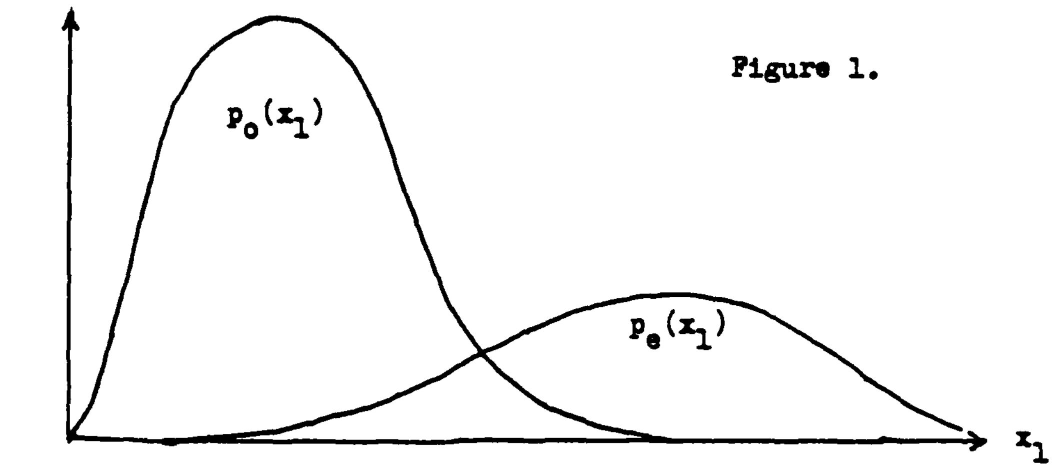


Figure 1.

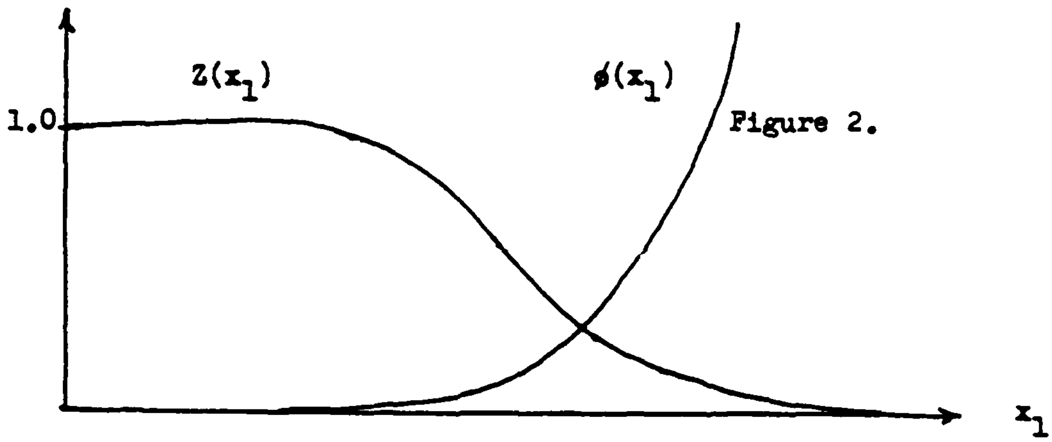


Figure 2.

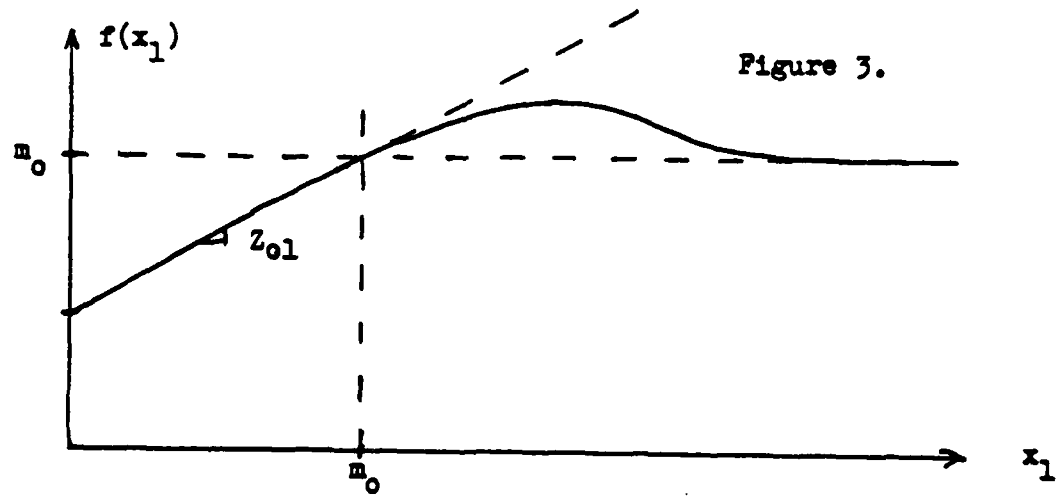


Figure 3.

very large, an excess observation is highly credible, and we settle for the "no-information" manual rating, m_o.

We see that the resulting forecast is quite similar to that obtained by ordinary credibility theory, but trimming the data and replacing x_1 by $\min(x_1,M)$ as in (1.2). Although sharp trimming will not have the "bump" shown in Figure 3, the effect will be small because the three parameters (a,b,M) in (1.2) can be adjusted to minimize the mean-squared error, thus giving a straight-line portion to the forecast which is slightly different than (4.2).

Another point in favor of the trimming is that it might be difficult to implement the exact predictive form in Figure 3 in a real experience-rating scheme; it would be difficult to explain a plan in which a policy with a larger claim might have a smaller next year's premium !

5. A Numerical Example

Figure 4 shows a numerical example in which normal densities were chosen for the average ordinary and excess densities; the means and standard deviations were:

$$m_o = \mu_o = 10 \ ; \ \sigma_o = 5 \ ; \qquad \mu_e = 50 \ ; \ \sigma_e = 20 \ .$$

A contamination probability $\pi = 0.1$ and a credibility factor $Z_{o1} = 0.5$ were used, so

$$f_o(x_1) = 10 + 0.5(x_1-10) \ .$$

The resulting exact forecast $f(x_1)$ is plotted in Figure 4, together with the optimal trimmed forecast,which is approximately

$$f(x_1) = 10 + 0.441\left[\min(x_1,14.7) - 10\right] \ .$$

Note that in the use of Gisler's results, one must subtract $\pi\mu_e$ from his forecast, as he does not have an explicit model for the generation of excess claims, and is predicting a future observation of either type.

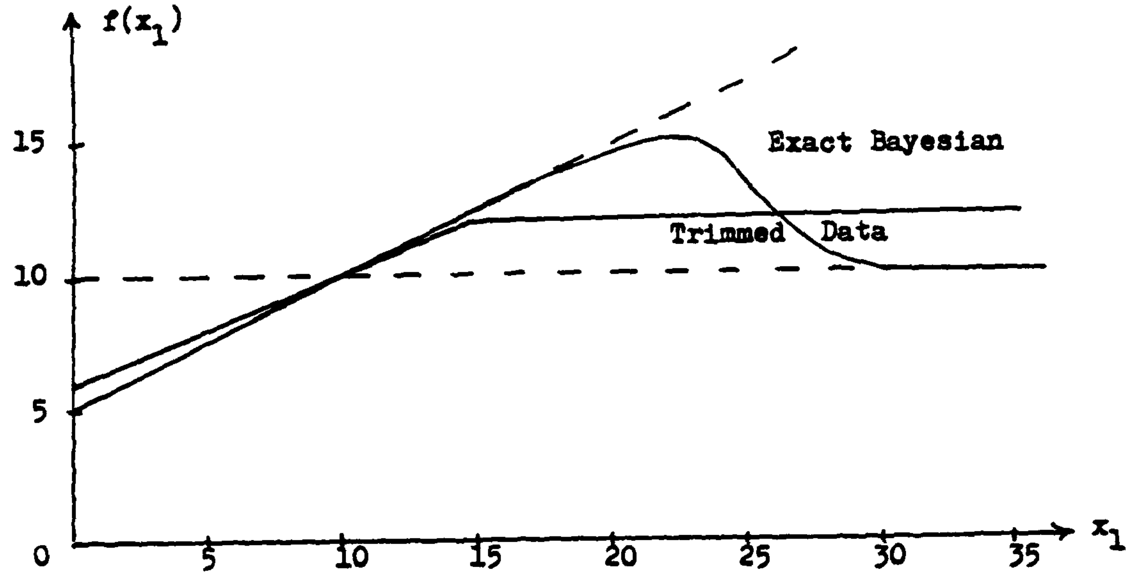


Figure 4. Exact Bayesian and Trimmed Data Forecasts for Example.

6. General Case

From the preceeding, it should be clear that the optimal predictor for an arbitrary number of observations n consists of 2^n terms from (2.2)(2.3), corresponding to all the different ways in which the data $\underline{x}=(x_1,x_2,\ldots x_n)$ can be partitioned into ordinary or excess categories. The formulae are greatly simplified in the general case if we use set-theoretic notation.

Let $\mathcal{N} = \{1,2,\ldots n\}$, $\mathcal{J}$ be any subset of $\mathcal{N}$ (including $\mathcal{N}$ and the empty set $\emptyset$), $\bar{\mathcal{J}} = \mathcal{N} - \mathcal{J}$, and $J = \|\mathcal{J}\|$. Use also $\mathcal{J}$ as a subscript to denote an arbitrary subset of observables, so that, for example, $x_{\mathcal{J}} = \{x_j | j \in \mathcal{J}\}$. Then the probability that this subset is all ordinary is:

$$p_o(x_{\mathcal{J}}) = \int \prod_{j\in\mathcal{J}} p_o(x_j|\theta)\, p(\theta)\, d\theta \quad , \tag{6.1}$$

whereas the probability that it is all excess is:

$$p_e(x_{\mathcal{J}}) = \prod_{j \in \mathcal{J}} p_e(x_j) \quad . \tag{6.2}$$

For consistency in the following equations, set $p_o(x_\emptyset) = p_e(x_\emptyset) = 1$.

Next, we define $f_o(x_{\mathcal{J}})$ to be the Bayesian forecast of an ordinary random variable, $\tilde{y}_o$, using only the data $x_{\mathcal{J}}$, assumed to be all of ordinary type. This might be the J-term generalization of (4.2), e.g., (2.1) with m replaced by m_o, Z_n replaced by Z_{oJ}, and of course using only the data $x_{\mathcal{J}}$. For consistency, the no-data forecast is $f_o(x_\emptyset) = m_o$.

Then, examination of the expansion of (2.3) shows that the forecast consists of the weighted sum of 2^n forecasts:

$$f(\underline{x}) = \sum_{\mathcal{J}} Z_{\mathcal{J}}(\underline{x}) \; f_o(x_{\mathcal{J}}) \quad , \tag{6.3}$$

where the data-dependent credibility factors are:

$$Z_{\mathcal{J}}(\underline{x}) = K \; (1-\pi)^J \pi^{n-J} \; p_o(x_{\mathcal{J}}) \; p_e(x_{\bar{\mathcal{J}}}) \quad , \tag{6.4}$$

and K is adjusted so that the factors sum to unity.

The sum in (6.3) is over all 2^n subsets of $\mathcal{N}$, although, as previously stated, it is unlikely that more than a few excess claims would be present with π small. This suggests the following computational strategy: arrange the data in decreasing magnitude, $x_1 > x_2 > \ldots > x_n$, take $\bar{\mathcal{J}}$ successively to be $\emptyset$, $\{1\}$, $\{1,2\}$, ...etc., and compute the corresponding credibility factors $Z_{\mathcal{J}}$. At some point these factors will become quite small, and the remaining terms in (6.3) can be neglected. One can, if desired, bound the neglected terms.

7. Continuing Research

The results presented here are part of a continuing research effort, joint with H. Bühlmann and A. Gisler. Current effort is devoted to multi-dimensional computations, and comparison of trimmed-data forecasts with the exact Bayesian prediction. Preliminary results indicate that the Gisler approximation continues to be quite good in the multi-dimensional case. These and other results will be be presented in an expanded version of this paper later this year.

References

1. A. Gisler, "Optimum Trimming of Data in the Credibility Model", Bulletin of the Society of Swiss Actuaries, 1980, Heft 3, 313-325 (1980).
2. R. Norberg, "The Credibility Approach to Experience Rating", Scandinavian Actuarial Journal, 1979, No. 4, 181-221 (1979).

FORSCHUNGSINSTITUT FUER MATHEMATIK ETH ZUERICH

Excess Claims and Data Trimming in the Context of Credibility Rating Procedures

by

Hans Bühlmann
ETH Zürich

Alois Gisler
Winterthur-Versicherungen, Winterthur

William S. Jewell
Forschungsinstitut für Mathematik ETH Zürich
and University of California, Berkeley

November 1981

Excess Claims and Data Trimming in the Context of Credibility Rating Procedures

by Hans Bühlmann, Alois Gisler, William S. Jewell*

1. Motivation

In Ratemaking and in Experience Rating one is often confronted with the dilemma of whether or not to fully charge very large claims to the claims load of small risk groups or of individual risks. Practitioners typically use an a posteriori argument in this situation: "If such large claims should be fully charged then the rates obtained would become 'ridiculous', hence it should not be done." The present paper aims at explaining this practical attitude from first principles.

Credibility Theory in its standard form makes the first step in the good direction. It explains to us that <u>all claims</u> should not be fully charged (but only with the constant fraction of the credibility weight). In many applications, however, it is still felt that the fraction of this charge should depend on the <u>size of a claim</u>. This leads very naturally to the idea of combining credibility procedures and data trimming.

Of course, such an idea needs to be tested. The first argument in favour of it was given by Gisler [1] who showed that in many cases the mean quadratic loss of the credibility estimator is substantially reduced if one introduces trimming of claims data. This paper goes even further. It formalizes the standard way of thinking about large claims and then shows that "optimal forecasting" of rates (using Bayes estimation techniques) and forecasting by "credibility techniques combined with data trimming" lead to almost identical results.

* The authors are greatly indebted to R. Schnieper who did all the numerical work on the ETH computer.

2. The Basic Model

Throughout the paper we work with the most simple model in the credibility context

$\underline{X} = (X_1, X_2, \ldots, X_n)$ is the random vector representing the experience of a given risk in the years $1, 2, \ldots, n$

- The quality of the risk is characterized by an unknown parameter value θ, which we consider as a realisation of a random variable Θ with distribution function $U(\theta)$
- Given the parameter value θ, $\{X_1, X_2, \ldots, X_n\}$ are i.i.d. with density function $f_\theta(x)$ [mean $\mu(\theta)$, variance $\sigma^2(\theta)$]

To these standard assumption in credibility theory we add now some more structure regarding the distribution of the size of a claim. The main idea is introduced by the assumption that the claims sizes are drawn from two different urns (distributions). Mostly, i.e. with probability $1-\pi$, we observe an <u>ordinary claim</u> with density $p_o(x/\theta)$ [mean $\mu_o(\theta)$, variance $\sigma_o^2(\theta)$] and occasionally, i.e. with probability π, we observe an <u>excess claim</u> (catastrophic claim) with density $p_e(x/\theta)$ [mean $\mu_e(\theta)$, variance $\sigma_e^2(\theta)$] .

	$p_o(x/\theta)$ ordinary claim amounts	$p_e(x/\theta)$ excess claim amounts
occurrence	$1-\pi$	π

We have assumed that the mixing probabilities are independent of θ and from now on we shall also suppose that the density of the excess claims is independent of the risk parameter, hence formalizing the idea that large catastrophic claims have no bearing on the quality of the risk.

In mathematical shorthand all the considerations just made regarding additional structure are summed up by stating that the density

$f_\theta(x)$ has the following form

1) $f_\theta(x) = (1-\pi)p_o(x/_\theta) + \pi p_e(x)$

3. The Basic Problem

As always in the credibility context our aim is to estimate

$\mu(\theta)$	based on the observations of	$\underline{X} = (X_1, X_2, \ldots, X_n)$
pure premium for the given risk		experience of the given risk in the years $1,2,\ldots,n$

One knows that the best estimator for this problem is

$$P[\underline{X}] = E[\mu(\theta)/_{\underline{X}}]$$

Using the special structure of formula 1) we obtain

2) $$P[\underline{X}] = \pi\mu_e + (1-\pi) \underbrace{E[\mu_o(\theta)/_{\underline{X}}]}_{g(\underline{X})}$$

If we use standard credibility techniques we estimate by

3) $$f[\underline{X}] = a + b\sum_{i=1}^{n} X_i \quad \text{with optimal choice of } a,b$$

And if in addition we introduce trimming of the data we estimate by

4) $$f[\underline{X}] = a + b\sum_{i=1}^{n} (X_i \wedge M) \text{ with optimal choice of } a,b,M$$

Using 4) we are committing the following error against optimal estimation

$$6)\quad \inf_{a,b,M} E\left\{P[\underline{X}]-f[\underline{X}]\right\}^2 = \inf_{a,b,M} E\left\{\pi\mu_e + (1-\pi)g(\underline{X}) - a - b\sum_{i=1}^{n}(X_i \wedge M)\right\}^2$$

$$= (1-\pi)^2 \inf_{a,b,M} E\left\{\frac{\pi\mu_e - a}{1-\pi} + g(\underline{X}) - \frac{b}{1-\pi}\sum_{i=1}^{n}(X_i \wedge M)\right\}^2$$

$$= (1-\pi)^2 \inf_{a',b',M} E\left\{g(\underline{X}) - a' - b'\sum_{i=1}^{n}(X_i \wedge M)\right\}^2$$

The following two problems are therefore equivalent

A) Estimate $P[\underline{X}]$ (total premium) by $a + b\sum_{i=1}^{n}(X_i \wedge M)$ with optimal a,b,M

B) Estimate $g(\underline{X})$ (ordinary premium) by $a' + b'\sum_{i=1}^{n}(X_i \wedge M)$ with optimal a',b',M

For the optimal choices of the parameters (denoted by ˜) we have

$$7)\quad \tilde{a} = (1-\pi)\,\tilde{a}' + \pi\mu_e$$

$$\tilde{b} = (1-\pi)\,\tilde{b}'$$

In the following we want to illustrate that $\tilde{a}' + \tilde{b}'\sum_{i=1}^{n}(X_i \wedge \tilde{M})$ is a good approximation of $g(\underline{X}) = E\,[\mu_o(\theta)/\underline{X}]$ (Problem B) above)

We actually shall compare

$\tilde{a}' + \tilde{b}'\sum(x_i \wedge \tilde{M})$ with $g(\underline{x})$ for any observation $\underline{x}$ of $\underline{X}$

4. The Exact Form of $g(\underline{x})$

Writing out the conditional expectation $E\left[\mu_o(\theta)/\underline{X}=\underline{x}\right]$ we obtain

$$8)\quad g(\underline{x}) = \frac{\int \mu_o(\theta)\left[\prod_{i=1}^{n}\left\{(1-\pi)p_o(x_{i/\theta}) + \pi p_e(x_i)\right\}\right] dU(\theta)}{\int\left[\prod_{i=1}^{n}\left\{(1-\pi)p_o(x_{i/\theta}) + \pi p_e(x_i)\right\}\right] dU(\theta)}$$

Putting $I = \{1,2,\ldots n\}$ and $S \subset I$ we rewrite

$$9)\quad \prod_{i=1}^{n}\left\{(1-\pi)p_o(x_{i/\theta})+\pi p_e(x_i)\right\} = \sum_{S\subset I}(1-\pi)^s\, \pi^{n-s} \prod_{i\in S} p_o(x_{i/\theta}) \prod_{i\in \bar{S}} p_e(x_i)$$

where the sum on the right side must be taken over all subsets

$S \subset I$ (including $\emptyset$ and I) with $s = |S|$
and $n = |I|$

We also use the abbreviations

$$p_o(x_S) = \int \prod_{i\in S} p_o(x_{i/\theta})\, dU(\theta)$$

$$p_e(x_S) = \int \prod_{i\in S} p_e(x_i)\, dU(\theta) = \prod_{i\in S} p_e(x_i)$$

$$L\,(x_S) = \left(\frac{\pi}{1-\pi}\right)^{n-s} \frac{p_o(x_S)\,p_e(x_{\bar{S}})}{p_o(x_I)} \qquad \left[p_o(x_\emptyset) = 1\right]$$

$$E_o\left[\mu_o(\theta)/x_S\right] = \frac{\int \mu_o(\theta) \prod_{i\in S} p_o(x_{i/\theta})\, dU(\theta)}{p_o(x_S)}$$

Then introducing 9) into 8) and carrying out the integration we find for the numerator of $g(\underline{x})$

$$\sum_{S\subset I}(1-\pi)^s\, \pi^{n-s} \prod_{i\in\bar{S}} p_e(x_i) \int \mu_o(\theta) \prod_{i\in S} p_o(x_{i/\theta})\, dU(\theta)$$

or

$$\sum_{S\subset I}(1-\pi)^s\, \pi^{n-s}\, p_o(x_S)\,p_e(x_{\bar{S}})\, E_o\left[\mu_o(\theta)/x_S\right]$$

and for the denominator of $g(\underline{x})$

$$\sum_{S\subset I} (1-\pi)^s \, \pi^{n-s} \, p_o(x_S) p_e(x_{\bar{S}})$$

Dividing both numerator and denominator by $(1-\pi)^n \, p_o(x_I)$ we finally arrive at

$$10) \quad \boxed{g(\underline{x}) = \frac{E_o\left[\mu_o(\theta)/\underline{x}\right] + \sum\limits_{\substack{S\subset I \\ S\neq I}} L(x_S) E_o\left[\mu_o(\theta)/x_S\right]}{1 + \sum\limits_{\substack{S\subset I \\ S\neq I}} L(x_S)}}$$

<u>Remarks:</u>

i) Observe that $g(\underline{x})$ is a weighted average of forecasts based on all subsamples x_S of the total sample x_I, the forecasts being calculated under the assumption that the subsample contains only claims of the ordinary type.

ii) As $\frac{\pi}{1-\pi}$ is usually rather small the weight of $L(x_S)$ is rather quick decreasing with decreasing number of observations in x_S; for a fixed number of observations the weight $L(x_S)$ is rather big if both $p_o(x_S)$ and $p_e(x_{\bar{S}})$ are big i.e. if x_S and $x_{\bar{S}}$ are very likely to come from the ordinary and the excess urn respectively.

iii) Dividing by $p_o(x_I)$ is obviously only allowed if all the observed claims are possibly of ordinary type. The weight function $L(x_S)$ is then only positive if $p_e(x_{\bar{S}})$ is positive i.e. if the subset $x_{\bar{S}}$ is possibly of excess type. Thus the formula does what we would have done by intuition as well, it excludes predictions based on claims which can be surely recognized as excess claims.

5. More insight from the single observation case

At this point it is worthwhile to consider the special case where the whole sample of observations contains only one observation, i.e.

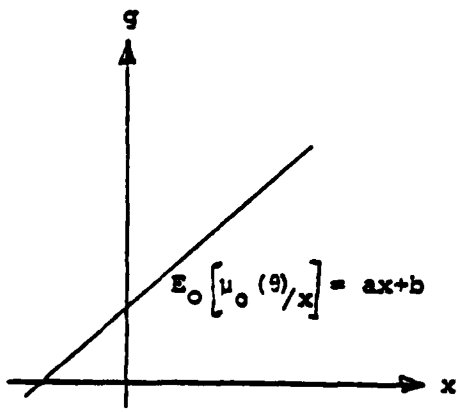
$\underline{x} = (x_1)$

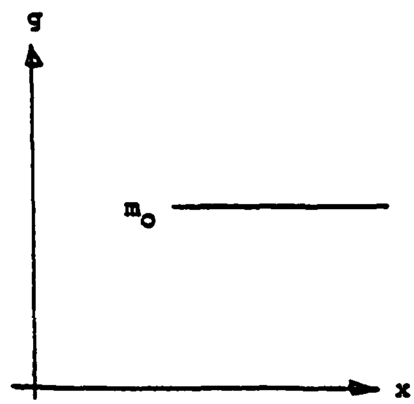
For simplicity we omit the index 1 and write x for the single observation. We have then

$$\text{11)} \quad g(x) = \frac{E_o\left[\mu_o(\theta)/x\right] + L\left[x_\emptyset\right] E\left[\mu_o(\theta)\right]}{1 + L\left[x_\emptyset\right]}$$

with $$L\left[x_\emptyset\right] = \frac{\pi}{1-\pi}\,\frac{p_e(x)}{p_o(x)}$$

The right hand side is a multiple of the likelihood ratio. If the latter is monotonically increasing (which is typically the case in applications) so is also the weight given to the constant estimator $E\left[\mu_o(\theta)\right] = m_o$. Assume in addition that $E_o\left[\mu_o(\theta)/x\right]$ is of linear form; then our estimator $g(x)$ is a mixture of the two cases (corresponding to the two pictures)





the weight being shifted from the estimator on the left to the estimator on the right as x increases. The resulting estimator is almost of the form $a+b \min(x,M)$. Hence credibility with trimming is almost exact! This fact will be illustrated by a numerical example in section 6. In fact our numerical example will show that this fact also carries over to higher dimensions.

6. A Numerical Example

6.1) For explicit calculations we are assuming that for ordinary claims

$p_o(x/\theta)$ is a normal density with mean θ, variance v

θ is normally distributed with mean m_o, variance w

We then have

$p_o(x_S) = \int \pi_{i\in S}\, p_o(x_i/\theta)\,dU(\theta)$ which turns out to be a multi-dimensional normal density with mean vector

$$\begin{pmatrix} m_o \\ m_o \\ \vdots \\ m_o \end{pmatrix}$$

and covariance matrix

$$\Sigma = \begin{pmatrix} w+v & w \ldots w \\ w & w+v \ldots \\ \vdots & \vdots \\ w & \ldots w+v \end{pmatrix}$$

hence

12) $$p_o(x_S) = \frac{\sqrt{|A|}}{(2\pi)^{s/2}}\, e^{-\frac{1}{2}\sum_{\substack{i\in S\\ j\in S}} a_{ij}(x_i-m_o)(x_j-m_o)}$$

with $A = \Sigma^{-1}$

Proof that $p_o(x_S)$ has density 12:

a) Given θ any linear combination $\sum_{i\in S} c_i x_i$ is normal with mean $\sum_{i\in S} c_i\theta$ and variance $\sum_{i\in S} c_i^2 v$. Integrated out with respect to the normal structure function of θ we obtain a normal distribution with mean $\sum_{i\in S} c_i m_o$ and variance $\left(\sum_{i\in S} c_i\right)^2 w + \sum_{i\in S} c_i^2 v$. But a sample x_S whose linear combinations are all normally distributed is multidimensional normal.

b) Let $\sum = \left(\sigma_{ij}\right)_{\substack{i\in S\\ j\in S}}$

$$\sigma_{ij} = \mathrm{Cov}(X_i,X_j) = E\left[\mathrm{Cov}(X_i,X_j)_{/\theta}\right] + \mathrm{Var}\left[E\left[X_{i/\theta}\right]\cdot E\left[X_{j/\theta}\right]\right]$$

$$= \delta_{ij}\, v + w$$

q.e.d.

It should be noted that

13) $\det \sum = v^n + nv^{n-1}w$ (subtract first row from all other rows and then develop along the first column)

Also observe the explicit form of

$\sum^{-1} = A = \left(a_{ij}\right)_{\substack{i\in S\\ j\in S}}$, namely

14) $a_{ij} = \frac{1}{v}\left(\delta_{ij} - \frac{w}{v+nw}\right)$ $\left[\text{use } (I+\alpha|\underline{\beta})^{-1} = I - \frac{\alpha|\underline{\beta}}{1+\underline{\alpha}\beta|}\right]$

From elementary calculations in credibility theory we finally also know that

15) $E_o\left[\theta/x_S\right] = \bar{x}_S \frac{sw}{v+sw} + m_o \frac{v}{v+sw}$

6.2) For the <u>excess claims</u> the probability law is specified by assuming that

$p_e(x)$ is a normal density with mean μ_e

variance σ_e^2

7. Numerical Calculations of $g(\underline{x})$

For our calculations we have chosen

$$m_o = 10 \qquad\qquad \mu_e = 50$$

$$\left.\begin{aligned} v &= 12.5 \\ w &= 12.5 \end{aligned}\right\} \sigma_o = 5 \qquad\qquad \sigma_e = 20$$

$$1-\pi = 0.9 \qquad\qquad \pi = 0.1$$

and we obtain

a) <u>for n=1 (single observation case)</u>

x	g(x)
5.	7.5091
6.	8.0068
7.	8.5049
8.	9.0033
9.	9.5017
10.	10.0000
11.	10.4979
12.	10.9951
13.	11.4910
14.	11.9850
15.	12.4755
16.	12.9602
17.	13.4348
18.	13.8919
19.	14.3177
20.	14.6876
21.	14.9586
22.	15.0602
23.	14.8946
24.	14.3712
25.	13.4968
26.	12.4521
27.	11.5065
28.	10.8265
29.	10.4153
30.	10.1952
31.	10.0870
32.	10.0370
33.	10.0151
34.	10.0059
35.	10.0022
36.	10.0008
37.	10.0003
38.	10.0001
39.	10.0000
40.	10.0000

b) <u>for n=2 (two observations)</u> $g(x_1,x_2)$

$x_1 \backslash x_2$	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	6.68	7.02	7.35	7.68	8.02	8.35	8.68	9.01	9.32	9.61	9.86	10.04	10.08	9.92	9.53	8.98	8.43	8.02	7.78	7.65
6	7.02	7.35	7.68	8.01	8.35	8.68	9.01	9.33	9.65	9.94	10.21	10.41	10.50	10.42	10.12	9.62	9.07	8.62	8.32	8.17
7	7.35	7.68	8.01	8.34	8.68	9.01	9.34	9.66	9.98	10.28	10.55	10.78	10.92	10.91	10.70	10.27	9.73	9.24	8.90	8.70
8	7.68	8.01	8.34	8.68	9.01	9.34	9.67	9.99	10.31	10.62	10.90	11.15	11.33	11.38	11.26	10.92	10.42	9.90	9.51	9.26
9	8.02	8.35	8.68	9.01	9.34	9.67	10.00	10.32	10.65	10.96	11.25	11.52	11.72	11.83	11.79	11.55	11.11	10.59	10.14	9.84
10	8.35	8.68	9.01	9.34	9.67	10.00	10.33	10.66	10.98	11.30	11.60	11.88	12.11	12.27	12.30	12.15	11.79	11.30	10.81	10.44
11	8.68	9.01	9.34	9.67	10.00	10.33	10.66	10.99	11.32	11.63	11.94	12.23	12.49	12.69	12.78	12.72	12.46	12.02	11.51	11.08
12	9.01	9.33	9.66	9.99	10.32	10.66	10.99	11.32	11.65	11.97	12.29	12.59	12.86	13.09	13.24	13.26	13.09	12.73	12.23	11.76
13	9.32	9.65	9.98	10.31	10.65	10.98	11.32	11.65	11.98	12.30	12.62	12.93	13.22	13.47	13.67	13.75	13.68	13.41	12.97	12.47
14	9.61	9.94	10.28	10.62	10.96	11.30	11.63	11.97	12.30	12.63	12.96	13.27	13.57	13.85	14.07	14.22	14.23	14.06	13.69	13.20
15	9.86	10.21	10.55	10.90	11.25	11.60	11.94	12.29	12.62	12.96	13.29	13.61	13.92	14.21	14.46	14.65	14.74	14.66	14.39	13.93
16	10.04	10.41	10.78	11.15	11.52	11.88	12.23	12.59	12.93	13.27	13.61	13.94	14.26	14.56	14.83	15.06	15.20	15.21	15.03	14.65
17	10.08	10.50	10.92	11.33	11.72	12.11	12.49	12.86	13.22	13.57	13.92	14.26	14.58	14.90	15.19	15.44	15.62	15.70	15.61	15.31
18	9.92	10.42	10.91	11.38	11.83	12.27	12.69	13.09	13.47	13.85	14.21	14.56	14.90	15.22	15.52	15.79	16.01	16.13	16.12	15.90
19	9.53	10.12	10.70	11.26	11.79	12.30	12.78	13.24	13.67	14.07	14.46	14.83	15.19	15.52	15.84	16.12	16.36	16.51	16.54	16.38
20	8.98	9.62	10.27	10.92	11.55	12.15	12.72	13.26	13.75	14.22	14.65	15.06	15.44	15.79	16.12	16.42	16.66	16.83	16.87	16.73
21	8.43	9.07	9.73	10.42	11.11	11.79	12.46	13.09	13.68	14.23	14.74	15.20	15.62	16.01	16.36	16.66	16.91	17.07	17.10	16.94
22	8.02	8.62	9.24	9.90	10.59	11.30	12.02	12.73	13.41	14.06	14.66	15.21	15.70	16.13	16.51	16.83	17.07	17.21	17.20	16.97
23	7.78	8.32	8.90	9.51	10.14	10.81	11.51	12.23	12.97	13.69	14.39	15.03	15.61	16.12	16.54	16.87	17.10	17.20	17.10	16.76
24	7.65	8.17	8.70	9.26	9.84	10.44	11.08	11.76	12.47	13.20	13.93	14.65	15.31	15.90	16.38	16.73	16.94	16.97	16.76	16.28

c) <u>for n=5 (five observations)</u>

$g(x_1,x_2, C_3,C_4,C_5)$ note: C_3,C_4,C_5 are chosen as "parameters" for the following tables

i) $(C_3,C_4,C_5) = (10,10,10)$

$x_1 \backslash x_2$	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	8.76	8.86	8.98	9.13	9.28	9.43	9.56	9.66	9.71	9.71	9.67	9.60	9.53	9.47	9.42	9.39	9.37	9.36	9.36	9.35
6	8.86	8.94	9.06	9.20	9.35	9.50	9.62	9.71	9.76	9.77	9.73	9.67	9.60	9.54	9.49	9.46	9.44	9.43	9.43	9.43
7	8.98	9.06	9.17	9.31	9.46	9.60	9.72	9.81	9.86	9.87	9.84	9.78	9.71	9.65	9.61	9.58	9.56	9.55	9.54	9.54
8	9.13	9.20	9.31	9.44	9.58	9.72	9.84	9.94	9.99	10.00	9.97	9.92	9.85	9.80	9.75	9.72	9.70	9.69	9.69	9.69
9	9.28	9.35	9.46	9.58	9.72	9.86	9.98	10.08	10.14	10.15	10.13	10.08	10.01	9.96	9.91	9.88	9.86	9.85	9.85	9.84
10	9.43	9.50	9.60	9.72	9.86	10.00	10.12	10.22	10.29	10.31	10.28	10.24	10.17	10.12	10.07	10.04	10.02	10.01	10.00	10.00
11	9.56	9.62	9.72	9.84	9.98	10.12	10.25	10.36	10.42	10.45	10.43	10.38	10.32	10.26	10.22	10.18	10.16	10.15	10.15	10.14
12	9.66	9.71	9.81	9.94	10.08	10.22	10.36	10.46	10.53	10.56	10.55	10.50	10.44	10.38	10.33	10.30	10.28	10.27	10.26	10.26
13	9.71	9.76	9.86	9.99	10.14	10.29	10.42	10.53	10.61	10.64	10.63	10.58	10.52	10.46	10.41	10.37	10.35	10.34	10.33	10.33
14	9.71	9.77	9.87	10.00	10.15	10.31	10.45	10.56	10.64	10.67	10.66	10.61	10.55	10.48	10.43	10.40	10.37	10.36	10.35	10.35
15	9.67	9.73	9.84	9.97	10.13	10.28	10.43	10.55	10.63	10.66	10.64	10.60	10.53	10.47	10.41	10.37	10.35	10.34	10.33	10.33
16	9.60	9.67	9.78	9.92	10.08	10.24	10.38	10.50	10.58	10.61	10.60	10.55	10.48	10.41	10.36	10.32	10.29	10.28	10.27	10.27
17	9.53	9.60	9.71	9.85	10.01	10.17	10.32	10.44	10.52	10.55	10.53	10.48	10.41	10.34	10.29	10.25	10.22	10.21	10.20	10.20
18	9.47	9.54	9.65	9.80	9.96	10.12	10.26	10.38	10.46	10.48	10.47	10.41	10.34	10.27	10.22	10.18	10.16	10.14	10.14	10.13
19	9.42	9.49	9.61	9.75	9.91	10.07	10.22	10.33	10.41	10.43	10.41	10.36	10.29	10.22	10.17	10.13	10.10	10.09	10.09	10.08
20	9.39	9.46	9.58	9.72	9.88	10.04	10.18	10.30	10.37	10.40	10.37	10.32	10.25	10.18	10.13	10.09	10.07	10.06	10.05	10.05
21	9.37	9.44	9.56	9.70	9.86	10.02	10.16	10.28	10.35	10.37	10.35	10.29	10.22	10.16	10.10	10.07	10.05	10.03	10.03	10.02
22	9.36	9.43	9.55	9.69	9.85	10.01	10.15	10.27	10.34	10.36	10.34	10.28	10.21	10.14	10.09	10.06	10.03	10.02	10.02	10.01
23	9.36	9.43	9.54	9.69	9.85	10.00	10.15	10.26	10.33	10.35	10.33	10.27	10.20	10.14	10.09	10.05	10.03	10.02	10.01	10.01
24	9.35	9.43	9.54	9.69	9.84	10.00	10.14	10.26	10.33	10.35	10.33	10.27	10.20	10.13	10.08	10.05	10.02	10.01	10.01	10.00

ii) $(C_3, C_4, C_5) = (10, 10, 25)$

$x_1 \backslash x_2$	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	8.59	8.70	8.84	9.01	9.18	9.35	9.50	9.61	9.66	9.66	9.62	9.54	9.46	9.39	9.33	9.30	9.27	9.26	9.26	9.25
6	8.70	8.80	8.93	9.09	9.26	9.43	9.57	9.67	9.73	9.73	9.69	9.62	9.54	9.47	9.41	9.38	9.36	9.35	9.34	9.34
7	8.84	8.93	9.06	9.22	9.38	9.54	9.68	9.78	9.84	9.85	9.81	9.74	9.67	9.60	9.54	9.51	9.49	9.48	9.47	9.47
8	9.01	9.09	9.22	9.37	9.53	9.68	9.82	9.93	9.99	10.00	9.97	9.90	9.83	9.76	9.71	9.68	9.66	9.65	9.64	9.64
9	9.18	9.26	9.38	9.53	9.69	9.84	9.98	10.09	10.16	10.17	10.15	10.09	10.01	9.95	9.90	9.86	9.84	9.83	9.82	9.82
10	9.35	9.43	9.54	9.68	9.84	10.00	10.14	10.26	10.33	10.35	10.33	10.27	10.20	10.13	10.08	10.05	10.02	10.01	10.01	10.00
11	9.50	9.57	9.68	9.82	9.98	10.14	10.29	10.41	10.48	10.51	10.49	10.44	10.37	10.30	10.25	10.21	10.19	10.18	10.17	10.17
12	9.61	9.67	9.78	9.93	10.09	10.26	10.41	10.53	10.61	10.64	10.63	10.57	10.51	10.44	10.38	10.34	10.32	10.31	10.30	10.30
13	9.66	9.73	9.84	9.99	10.16	10.33	10.48	10.61	10.70	10.73	10.72	10.67	10.60	10.53	10.47	10.43	10.40	10.39	10.38	10.38
14	9.66	9.73	9.85	10.00	10.17	10.35	10.51	10.64	10.73	10.77	10.75	10.70	10.63	10.56	10.50	10.46	10.43	10.42	10.41	10.41
15	9.62	9.69	9.81	9.97	10.15	10.33	10.49	10.63	10.72	10.75	10.74	10.69	10.61	10.54	10.48	10.43	10.41	10.39	10.38	10.38
16	9.54	9.62	9.74	9.90	10.09	10.27	10.44	10.57	10.67	10.70	10.69	10.63	10.55	10.47	10.41	10.37	10.34	10.33	10.32	10.31
17	9.46	9.54	9.67	9.83	10.01	10.20	10.37	10.51	10.60	10.63	10.61	10.55	10.47	10.39	10.33	10.29	10.26	10.24	10.24	10.23
18	9.39	9.47	9.60	9.76	9.95	10.13	10.30	10.44	10.53	10.56	10.54	10.47	10.39	10.32	10.25	10.21	10.18	10.17	10.16	10.16
19	9.33	9.41	9.54	9.71	9.90	10.08	10.25	10.38	10.47	10.50	10.48	10.41	10.33	10.25	10.19	10.15	10.12	10.11	10.10	10.10
20	9.30	9.38	9.51	9.68	9.86	10.05	10.21	10.34	10.43	10.46	10.43	10.37	10.29	10.21	10.15	10.11	10.08	10.07	10.06	10.05
21	9.27	9.36	9.49	9.66	9.84	10.02	10.19	10.32	10.40	10.43	10.41	10.34	10.26	10.18	10.12	10.08	10.05	10.04	10.03	10.03
22	9.26	9.35	9.48	9.65	9.83	10.01	10.18	10.31	10.39	10.42	10.39	10.33	10.24	10.17	10.11	10.07	10.04	10.03	10.02	10.02
23	9.26	9.34	9.47	9.64	9.82	10.01	10.17	10.30	10.38	10.41	10.38	10.32	10.24	10.16	10.10	10.06	10.03	10.02	10.01	10.01
24	9.25	9.34	9.47	9.64	9.82	10.00	10.17	10.30	10.38	10.41	10.38	10.31	10.23	10.16	10.10	10.05	10.03	10.02	10.01	10.01

8. Optimal Trimming

Gisler has shown [1] that for given M the optimal choice of the approximation

$$\widehat{\mu(\theta)} = a + b \sum_{i=1}^{n} (x_i \wedge M) \quad \text{to} \quad \mu(\theta) \quad \left[\text{and hence to} \quad P[\underline{x}]\right]$$

can be calculated as follows

$$16) \quad \tilde{b} = \frac{b_1}{(n-1)b_2 + b_3} \qquad \text{where} \qquad \begin{aligned} b_1 &= \mathrm{Cov}\,[X_1 \wedge M,\ X_2] \\ b_2 &= \mathrm{Cov}\,[X_1 \wedge M,\ X_2 \wedge M] \\ b_3 &= \mathrm{Var}\,[X_1 \wedge M] \end{aligned}$$

$$17) \quad \tilde{a} + n\tilde{b}\, E[X \wedge M] = E[X]$$

With this optimal choice we then have

$$18) \quad E\left[\left(\widehat{\mu(\theta)} - \mu(\theta)\right)^2\right] = w - n\,\tilde{b} \cdot b_1$$

Hence the trimming point M is optimal if $\tilde{b} \cdot b_1$ is maximum.

In our basic model (cf. section 2) we find

19) $b_1 = (1-\pi)^2 \operatorname{Cov}\left[\mu_o^M(\theta),\ \mu_o(\theta)\right]$ where

$$\mu_o^M(\theta) = E\left[X\wedge M/\theta, X \text{ ordinary}\right]$$

$$\mu_o(\theta) = E\left[X/\theta, X \text{ ordinary}\right]$$

$$b_2 = (1-\pi)^2 \operatorname{Var}\left[\mu_o^M(\theta)\right]$$

$$b_3 = (1-\pi)E\left[\sigma_o^{2M}(\theta)\right] + \pi\sigma_e^{2M} + (1-\pi)^2 \operatorname{Var}\left[\mu_o^M(\theta)\right] + \pi(1-\pi)\ E\left[\left(\mu_o(\theta) - \mu_e^M\right)^2\right]$$

with $\sigma_o^{2M}(\theta) = \operatorname{Var}\left[X\wedge M/\theta, X \text{ ordinary}\right]$

$$\sigma_e^{2M} = \operatorname{Var}\left[X\wedge M/X \text{ excess}\right]$$

$$\mu_e^M = E\left[X\wedge M/X \text{ excess}\right]$$

Using explicitely the normal distribution as assumed both for ordinary and excess claims in section 6 we obtain from some rather tedious integrations:

Let $\Phi(.)$ denote the standardized normal distribution function and $\varphi(.)$ the standardized normal density function, then

20) $b_1 = (1-\pi)^2 w\Phi\left(\dfrac{M-m_o}{\sigma_o}\right)$ $\qquad \sigma_o = \sqrt{v + w}$

$$b_2 = (1-\pi)^2 \operatorname{Cov}\left[U_1\wedge M\ ,\ U_2\wedge M\right]$$

where the covariance is obtained by numerical integration.

Notation: (U_1,U_2) is $N\left(\begin{matrix} m_o \\ m_o \end{matrix}\ ,\ \Sigma\right)$ with $\Sigma = \begin{pmatrix} v+w & w \\ w & v+w \end{pmatrix}$

$$b_3 = A - B^2$$

$$A = (1-\pi)\left[(m_o^2+\sigma_o^2)\ \Phi\left(\frac{M-m_o}{\sigma_o}\right) - \sigma_o(M+m\ )\ \varphi\left(\frac{M-m_o}{\sigma_o}\right)\right]$$

$$+\ \pi\ \left[(\mu_e^2+\sigma_e^2)\ \Phi\left(\frac{M-\mu_e}{\sigma_e}\right) - \sigma_e(M+\mu_e)\ \varphi\left(\frac{M-\mu_e}{\sigma_e}\right)\right]$$

$$+\ M^2\ \left[1-(1-\pi)\Phi\left(\frac{M-m_o}{\sigma_o}\right) - \pi\Phi\left(\frac{M-\mu_e}{\sigma_e}\right)\right]$$

$$B = (1-\pi)\ \left[(m_o-M)\ \Phi\left(\frac{M-m_o}{\sigma_o}\right) - \sigma_o\varphi\left(\frac{M-m_o}{\sigma_o}\right)\right]$$

$$+\ \pi\ \left[(\mu_e-M)\ \Phi\left(\frac{M-\mu_e}{\sigma_e}\right) - \sigma_e\varphi\left(\frac{M-\mu_e}{\sigma_e}\right)\right]$$

$$+\ M$$

9. Numerical Calculations of $\tilde{a} + \tilde{b} \sum_{i=1}^{n} (x_i \wedge \tilde{M})$

Using the same parameter values as in section 7 we obtain the forecasts based on optimal trimming. To compare with $g(\underline{x})$ it is worthwhile to calculate also $\tilde{a}' + \tilde{b}' \sum_{i=1}^{n} (x_i \wedge \tilde{M})$
with

$$\tilde{a}' = \frac{\tilde{a} - \pi\mu_e}{1-\pi} \qquad\qquad \tilde{b}' = \frac{\tilde{b}}{1-\pi}$$

a) Results for n=1 (single observation case)

Truncation point $\tilde{M} = 14.68$

formula: $\hat{P}=0.4412(x \wedge \tilde{M})+9.5817$ $\hat{g}=0.4902(x \wedge \tilde{M})+5.0908$

x		
5	11.79	7.54
6	12.23	8.03
7	12.67	8.52
8	13.11	9.01
9	13.55	9.50
10	13.99	9.99
11	14.43	10.48
12	14.88	10.97
13	15.32	11.46
14	15.76	11.95
15	16.06	12.29
16	16.06	12.29
17	16.06	12.29
18	16.06	12.29
19	16.06	12.29
20	16.06	12.29

b) Results for n=2

truncation point $\tilde{M} = 19.52$

b_1) approximation to total premium $P[\underline{x}]$

formula: $$\hat{P} = 0.2289 \sum_{i=1}^{2} (x_i \wedge \tilde{M}) + 9.0351$$

$x_1 \backslash x_2$	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	11.32	11.55	11.78	12.01	12.24	12.47	12.70	12.93	13.16	13.38	13.61	13.84	14.07	14.30	14.53	14.65	14.65	14.65	14.65	14.65
6	11.55	11.78	12.01	12.24	12.47	12.70	12.93	13.16	13.38	13.61	13.84	14.07	14.30	14.53	14.76	14.88	14.88	14.88	14.88	14.88
7	11.78	12.01	12.24	12.47	12.70	12.93	13.16	13.38	13.61	13.84	14.07	14.30	14.53	14.76	14.99	15.11	15.11	15.11	15.11	15.11
8	12.01	12.24	12.47	12.70	12.93	13.16	13.38	13.61	13.84	14.07	14.30	14.53	14.76	14.99	15.22	15.33	15.33	15.33	15.33	15.33
9	12.24	12.47	12.70	12.93	13.16	13.38	13.61	13.84	14.07	14.30	14.53	14.76	14.99	15.22	15.44	15.56	15.56	15.56	15.56	15.56
10	12.47	12.70	12.93	13.16	13.38	13.61	13.84	14.07	14.30	14.53	14.76	14.99	15.22	15.44	15.67	15.79	15.79	15.79	15.79	15.79
11	12.70	12.93	13.16	13.38	13.61	13.84	14.07	14.30	14.53	14.76	14.99	15.22	15.44	15.67	15.90	16.02	16.02	16.02	16.02	16.02
12	12.93	13.16	13.38	13.61	13.84	14.07	14.30	14.53	14.76	14.99	15.22	15.44	15.67	15.90	16.13	16.25	16.25	16.25	16.25	16.25
13	13.16	13.38	13.61	13.84	14.07	14.30	14.53	14.76	14.99	15.22	15.44	15.67	15.90	16.13	16.36	16.48	16.48	16.48	16.48	16.48
14	13.38	13.61	13.84	14.07	14.30	14.53	14.76	14.99	15.22	15.44	15.67	15.90	16.13	16.36	16.59	16.71	16.71	16.71	16.71	16.71
15	13.61	13.84	14.07	14.30	14.53	14.76	14.99	15.22	15.44	15.67	15.90	16.13	16.36	16.59	16.82	16.94	16.94	16.94	16.94	16.94
16	13.84	14.07	14.30	14.53	14.76	14.99	15.22	15.44	15.67	15.90	16.13	16.36	16.59	16.82	17.05	17.17	17.17	17.17	17.17	17.17
17	14.07	14.30	14.53	14.76	14.99	15.22	15.44	15.67	15.90	16.13	16.36	16.59	16.82	17.05	17.28	17.39	17.39	17.39	17.39	17.39
18	14.30	14.53	14.76	14.99	15.22	15.44	15.67	15.90	16.13	16.36	16.59	16.82	17.05	17.28	17.50	17.62	17.62	17.62	17.62	17.62
19	14.53	14.76	14.99	15.22	15.44	15.67	15.90	16.13	16.36	16.59	16.82	17.05	17.28	17.50	17.73	17.85	17.85	17.85	17.85	17.85
20	14.65	14.88	15.11	15.33	15.56	15.79	16.02	16.25	16.48	16.71	16.94	17.17	17.39	17.62	17.85	17.97	17.97	17.97	17.97	17.97
21	14.65	14.88	15.11	15.33	15.56	15.79	16.02	16.25	16.48	16.71	16.94	17.17	17.39	17.62	17.85	17.97	17.97	17.97	17.97	17.97
22	14.65	14.88	15.11	15.33	15.56	15.79	16.02	16.25	16.48	16.71	16.94	17.17	17.39	17.62	17.85	17.97	17.97	17.97	17.97	17.97
23	14.65	14.88	15.11	15.33	15.56	15.79	16.02	16.25	16.48	16.71	16.94	17.17	17.39	17.62	17.85	17.97	17.97	17.97	17.97	17.97
24	14.65	14.88	15.11	15.33	15.56	15.79	16.02	16.25	16.48	16.71	16.94	17.17	17.39	17.62	17.85	17.97	17.97	17.97	17.97	17.97

b_2) approximation to ordinary premium $q(\underline{x})$

formula: $\hat{q} = 0.2543 \sum_{i=1}^{2} (x_i \wedge \tilde{M}) + 4.4834$

$x_1 \backslash x_2$	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	7.03	7.28	7.53	7.79	8.04	8.30	8.55	8.81	9.06	9.32	9.57	9.82	10.08	10.33	10.59	10.72	10.72	10.72	10.72	10.72
6	7.28	7.53	7.79	8.04	8.30	8.55	8.81	9.06	9.32	9.57	9.82	10.08	10.33	10.59	10.84	10.97	10.97	10.97	10.97	10.97
7	7.53	7.79	8.04	8.30	8.55	8.81	9.06	9.32	9.57	9.82	10.08	10.33	10.59	10.84	11.10	11.23	11.23	11.23	11.23	11.23
8	7.79	8.04	8.30	8.55	8.81	9.06	9.32	9.57	9.82	10.08	10.33	10.59	10.84	11.10	11.35	11.48	11.48	11.48	11.48	11.48
9	8.04	8.30	8.55	8.81	9.06	9.32	9.57	9.82	10.08	10.33	10.59	10.84	11.10	11.35	11.60	11.74	11.74	11.74	11.74	11.74
10	8.30	8.55	8.81	9.06	9.32	9.57	9.82	10.08	10.33	10.59	10.84	11.10	11.35	11.60	11.86	11.99	11.99	11.99	11.99	11.99
11	8.55	8.81	9.06	9.32	9.57	9.82	10.08	10.33	10.59	10.84	11.10	11.35	11.60	11.86	12.11	12.24	12.24	12.24	12.24	12.24
12	8.81	9.06	9.32	9.57	9.82	10.08	10.33	10.59	10.84	11.10	11.35	11.60	11.86	12.11	12.37	12.50	12.50	12.50	12.50	12.50
13	9.06	9.32	9.57	9.82	10.08	10.33	10.59	10.84	11.10	11.35	11.60	11.86	12.11	12.37	12.62	12.75	12.75	12.75	12.75	12.75
14	9.32	9.57	9.82	10.08	10.33	10.59	10.84	11.10	11.35	11.60	11.86	12.11	12.37	12.62	12.88	13.01	13.01	13.01	13.01	13.01
15	9.57	9.82	10.08	10.33	10.59	10.84	11.10	11.35	11.60	11.86	12.11	12.37	12.62	12.88	13.13	13.26	13.26	13.26	13.26	13.26
16	9.82	10.08	10.33	10.59	10.84	11.10	11.35	11.60	11.86	12.11	12.37	12.62	12.88	13.13	13.38	13.52	13.52	13.52	13.52	13.52
17	10.08	10.33	10.59	10.84	11.10	11.35	11.60	11.86	12.11	12.37	12.62	12.88	13.13	13.38	13.64	13.77	13.77	13.77	13.77	13.77
18	10.33	10.59	10.84	11.10	11.35	11.60	11.86	12.11	12.37	12.62	12.88	13.13	13.38	13.64	13.89	14.02	14.02	14.02	14.02	14.02
19	10.59	10.84	11.10	11.35	11.60	11.86	12.11	12.37	12.62	12.88	13.13	13.38	13.64	13.89	14.15	14.28	14.28	14.28	14.28	14.28
20	10.72	10.97	11.23	11.48	11.74	11.99	12.24	12.50	12.75	13.01	13.26	13.52	13.77	14.02	14.28	14.41	14.41	14.41	14.41	14.41
21	10.72	10.97	11.23	11.48	11.74	11.99	12.24	12.50	12.75	13.01	13.26	13.52	13.77	14.02	14.28	14.41	14.41	14.41	14.41	14.41
22	10.72	10.97	11.23	11.48	11.74	11.99	12.24	12.50	12.75	13.01	13.26	13.52	13.77	14.02	14.28	14.41	14.41	14.41	14.41	14.41
23	10.72	10.97	11.23	11.48	11.74	11.99	12.24	12.50	12.75	13.01	13.26	13.52	13.77	14.02	14.28	14.41	14.41	14.41	14.41	14.41
24	10.72	10.97	11.23	11.48	11.74	11.99	12.24	12.50	12.75	13.01	13.26	13.52	13.77	14.02	14.28	14.41	14.41	14.41	14.41	14.41

c) Results for n=5

truncation point $\tilde{M} = 22.83$

formulae: $\hat{P} = 0.1241 \sum_{i=1}^{5} (x_i \wedge \tilde{M}) + 7.0561$ total premium

$\hat{q} = 0.1378 \sum_{i=1}^{5} (x_i \wedge \tilde{M}) + 2.2845$ ordinary premium

c_1) approximation to total premium $P[\underline{x}]$ $= P[x_1, x_2, \underbrace{C_3, C_4, C_5}]$ chosen as fixed parameter values

i) $(C_3, C_4, C_5) = (10,10,10)$

$x_1 \backslash x_2$	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	12.02	12.14	12.27	12.39	12.52	12.64	12.76	12.89	13.01	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.23	14.23
6	12.14	12.27	12.39	12.52	12.64	12.76	12.89	13.01	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.36	14.36
7	12.27	12.39	12.52	12.64	12.76	12.89	13.01	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.48	14.48
8	12.39	12.52	12.64	12.76	12.89	13.01	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.61	14.61
9	12.52	12.64	12.76	12.89	13.01	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.73	14.73
10	12.64	12.76	12.89	13.01	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.85	14.85
11	12.76	12.89	13.01	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	14.98	14.98
12	12.89	13.01	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.10	15.10
13	13.01	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.23	15.23
14	13.14	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.25	15.35	15.35
15	13.26	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.25	15.37	15.47	15.47
16	13.39	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.25	15.37	15.49	15.60	15.60
17	13.51	13.63	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.29	15.37	15.49	15.62	15.72	15.72
18	13.63	13.76	13.86	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.25	15.37	15.49	15.62	15.74	15.85	15.85
19	13.76	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.25	15.37	15.49	15.62	15.74	15.87	15.97	15.97
20	13.88	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.25	15.37	15.49	15.62	15.74	15.87	15.99	16.09	16.09
21	14.01	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.25	15.37	15.49	15.62	15.74	15.87	15.99	16.12	16.22	16.22
22	14.13	14.25	14.38	14.50	14.63	14.75	14.87	15.00	15.12	15.25	15.37	15.49	15.62	15.74	15.87	15.99	16.12	16.24	16.34	16.34
23	14.23	14.36	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.45	16.45
24	14.23	14.36	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.45	16.45

ii) $(C_3,C_4,C_5) = (10,10,25)$

$x_1 \backslash x_2$	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	13.61	13.74	13.86	13.98	14.11	14.23	14.36	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.83	15.83
6	13.74	13.86	13.98	14.11	14.23	14.36	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.95	15.95
7	13.86	13.98	14.11	14.23	14.36	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.07	16.07
8	13.98	14.11	14.23	14.36	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.20	16.20
9	14.11	14.23	14.36	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.32	16.32
10	14.23	14.36	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.45	16.45
11	14.36	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.57	16.57
12	14.48	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.69	16.69
13	14.61	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	16.82	16.82
14	14.73	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	16.84	16.94	16.94
15	14.85	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	16.84	16.96	17.07	17.07
16	14.98	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	16.84	16.96	17.09	17.19	17.19
17	15.10	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	16.84	16.96	17.09	17.21	17.31	17.31
18	15.23	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	16.84	16.96	17.09	17.21	17.34	17.44	17.44
19	15.35	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	16.84	16.96	17.09	17.21	17.34	17.46	17.56	17.56
20	15.47	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	15.84	16.96	17.09	17.21	17.34	17.46	17.58	17.69	17.69
21	15.60	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	16.84	16.96	17.09	17.21	17.34	17.46	17.58	17.71	17.81	17.81
22	15.72	15.85	15.97	16.09	16.22	16.34	16.47	16.59	16.71	16.84	16.96	17.09	17.21	17.34	17.46	17.58	17.71	17.83	17.93	17.93
23	15.83	15.95	16.07	16.20	16.32	16.45	16.57	16.69	16.82	16.94	17.07	17.19	17.31	17.44	17.56	17.69	17.81	17.93	18.04	18.04
24	15.83	15.95	16.07	16.20	16.32	16.45	16.57	16.69	16.82	16.94	17.07	17.19	17.31	17.44	17.56	17.69	17.81	17.93	18.04	18.04

$\underline{c_2)\ \text{approximation to ordinary premium}}$ $g(\underline{x}) = g(x_1, x_2, \underbrace{C_1, C_2, C_3}_{\text{chosen as fixed parameters}})$

i) $(C_3,C_4,C_5) = (10,10,10)$

$x_1 \backslash x_2$	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	7.80	7.93	8.07	8.21	8.35	8.49	8.62	8.76	8.90	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.25	10.25
6	7.93	8.07	8.21	8.35	8.49	8.62	8.76	8.90	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.39	10.39
7	8.07	8.21	8.35	8.49	8.62	8.76	8.90	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.53	10.53
8	8.21	8.35	8.49	8.62	8.76	8.90	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.67	10.67
9	8.35	8.49	8.62	8.76	8.90	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.80	10.80
10	8.49	8.62	8.76	8.90	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.94	10.94
11	8.62	8.76	8.90	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.08	11.08
12	8.76	8.90	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.22	11.22
13	8.90	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.24	11.36	11.36
14	9.04	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.24	11.38	11.49	11.49
15	9.17	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.24	11.38	11.52	11.63	11.63
16	9.31	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.24	11.38	11.52	11.65	11.77	11.77
17	9.45	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.24	11.38	11.52	11.65	11.79	11.91	11.91
18	9.59	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.24	11.38	11.52	11.65	11.79	11.93	12.04	12.04
19	9.73	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.24	11.38	11.52	11.65	11.79	11.93	12.07	12.18	12.18
20	9.86	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.24	11.38	11.52	11.65	11.79	11.93	12.07	12.21	12.32	12.32
21	10.00	10.14	10.28	10.41	10.55	10.69	10.83	10.97	11.10	11.24	11.38	11.52	11.65	11.79	11.93	12.07	12.21	12.34	12.46	12.46
22	10.14	10.28	10.41	10.53	10.69	10.83	10.97	11.10	11.24	11.38	11.52	11.65	11.79	11.93	12.07	12.21	12.34	12.48	12.60	12.60
23	10.25	10.39	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.71	12.71
24	10.25	10.39	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.71	12.71

ii) $(C_3, C_4, C_5) = (10,10,25)$

$z_1 \backslash z_2$	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	9.56	9.70	9.84	9.98	10.12	10.25	10.39	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.02	12.02
6	9.70	9.84	9.98	10.12	10.25	10.39	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.16	12.16
7	9.84	9.98	10.12	10.25	10.39	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.30	12.30
8	9.98	10.12	10.25	10.39	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.43	12.43
9	10.12	10.25	10.39	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.57	12.57
10	10.25	10.39	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.71	12.71
11	10.39	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.85	12.85
12	10.53	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	12.99	12.99
13	10.67	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.12	13.12
14	10.80	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.15	13.26	13.26
15	10.94	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.15	13.29	13.40	13.40
16	11.08	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.15	13.29	13.42	13.54	13.54
17	11.22	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.15	13.29	13.42	13.56	13.68	13.68
18	11.36	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.15	13.29	13.42	13.56	13.70	13.81	13.81
19	11.49	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.15	13.29	13.42	13.56	13.70	13.84	13.95	13.95
20	11.63	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.15	13.29	13.42	13.56	13.70	13.84	13.97	14.09	14.09
21	11.77	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.15	13.29	13.42	13.56	13.70	13.84	13.97	14.11	14.23	14.23
22	11.91	12.04	12.18	12.32	12.46	12.60	12.73	12.87	13.01	13.15	13.29	13.42	13.56	13.70	13.84	13.97	14.11	14.25	14.36	14.36
23	12.02	12.16	12.30	12.43	12.57	12.71	12.85	12.99	13.12	13.26	13.40	13.54	13.68	13.81	13.95	14.09	14.23	14.36	14.48	14.48
24	12.02	12.16	12.30	12.43	12.57	12.71	12.85	12.99	13.12	13.26	13.40	13.54	13.68	13.81	13.95	14.09	14.23	14.36	14.48	14.48

10. Final Remarks

The Data Trimmed Credibility Formulae seem quite appropriate for Experience Rating in the presence of catastrophic (or as called in this paper excess) claims. With this intuitive background in our minds we have in our explicit calculations been looking at deviations from ordinary claims towards the higher side only. Obviously the normal distribution being symmetric one could also observe "outliers" to ordinary claims towards the lower side hence leading to a truncation at the lower end as well. But of course our assumption of normally distributed claims should only be seen as an approximation to the real world, and it is our feeling that the approximation is particularly bad at the lower tail of the distribution.

In any case truncation at the upper end of the distribution is introducing an additional parameter into the credibility formulae and we hope to have demonstrated in this paper that the labour caused by the new parameter can be worthwhile indeed.

11. Bibliography

[1] A. Gisler Optimales Stutzen von Beobachtungen im Credibility Modell (ETH Thesis 1980)

see also A. Gisler Optimum Trimming of Data in the Credibility Model BASA 1980 (3)

12. Appendix

For the interested reader we are attaching the explicit calculations leading to formulae 19) and 20).

A: Calculations leading to formula 19)

$$b_1 = E\left[Cov[X_1 \wedge M, X_{2/\theta}]\right] + Cov\left[E[X_1 \wedge M_{/\theta}], E[X_{2/\theta}]\right]$$

$Cov[X_1 \wedge M, X_{2/\theta}] = 0$, because X_1, X_2 are conditionally independent.

Hence

$$b_1 = Cov[(1-\pi)\mu_o^M(\theta) + \pi\mu_e^M, (1-\pi)\mu_o(\theta) + \pi\mu_e], \text{ or}$$

$$\boxed{b_1 = (1-\pi)^2 Cov[\mu_o^M(\theta), \mu_o(\theta)]}$$

and analogously (with $X_2 \wedge M$ instead of X_2)

$$\boxed{b_2 = (1-\pi)^2 Var[\mu_o^M(\theta)]} .$$

Let be $Y = 1_A$ where A denotes the event {X is ordinary}. Then

$$Var[X \wedge M_{/\theta}] = E\left[Var[X \wedge M_{/\theta,Y}]\right] + Var\left[E[X \wedge M_{/\theta,Y}]\right]$$

$$= (1-\pi)\sigma_o^{2M}(\theta) + \pi\sigma_e^{2M} + \pi(1-\pi)(\mu_o^M(\theta) - \mu_e^M)^2$$

Hence

$$b_3 = Var[X \wedge M]$$

$$= E\left[Var[X \wedge M_{/\theta}]\right] + Var[(1-\pi)\mu_o^M(\theta) + \mu_e^M], \text{ or}$$

$$\boxed{b_3 = (1-\pi) E[\sigma_o^{2M}(\theta)] + \pi\sigma_e^{2M} + \pi(1-\pi) E[(\mu_o^M(\theta) - \mu_e^M)^2] + (1-\pi)^2 Var\ \mu_o^M(\theta)]}$$

B: Calculations leading to formula 20)

i) Preparations

In the following we put $r=\sqrt{v}$, $s=\sqrt{w}$ and $\sigma_o=\sqrt{v+w}=\sqrt{r^2+s^2}$. Furthermore we denote by $\Phi(x)$ the standardized normal distribution function and by $\varphi(x)$ the standardized normal density function.

By convolution we get

$$\int_{-\infty}^{\infty} \frac{1}{s}\,\varphi\left(\frac{x-\mu}{s}\right)\Phi\left(\frac{M-x}{r}\right)dx = \Phi\left(\frac{M-\mu}{\sigma_o}\right)$$

$$\int_{-\infty}^{\infty} \frac{1}{rs}\,\varphi\left(\frac{x-\mu}{s}\right)\varphi\left(\frac{M-x}{r}\right)dx = \frac{1}{\sigma_o}\,\varphi\left(\frac{M-\mu}{\sigma_o}\right)$$

Noting that $\varphi'(x) = -x\,\varphi(x)$ integration by parts gives

$$\int_{-\infty}^{\infty} (x-\mu)\,\varphi\left(\frac{x-\mu}{s}\right)\Phi\left(\frac{M-x}{r}\right)dx = -\frac{s^2}{r}\int_{-\infty}^{\infty} \varphi\left(\frac{x-\mu}{s}\right)\varphi\left(\frac{M-x}{r}\right)dx$$

and thus

$$\int_{-\infty}^{\infty} x\varphi\left(\frac{x-\mu}{s}\right)\Phi\left(\frac{M-x}{r}\right)dx = \mu s\Phi\left(\frac{M-\mu}{\sigma_o}\right) - \frac{s^3}{\sigma_o}\,\varphi\left(\frac{M-\mu}{\sigma_o}\right)$$

Because of $\varphi\left(\frac{x-\mu}{s}\right)\varphi\left(\frac{M-x}{r}\right) = \varphi\left(\frac{M-\mu}{\sigma_o}\right)\varphi\left(\frac{x-\tilde{\mu}}{\tilde{\sigma}}\right)$

where $\tilde{\mu} = \dfrac{r^2\mu + s^2M}{r^2+s^2}$ and $\tilde{\sigma} = \dfrac{rs}{\sigma_o}$

we obtain

$$\int_{-\infty}^{\infty} x\varphi\left(\frac{x-\mu}{s}\right)\varphi\left(\frac{M-x}{r}\right)dx = \varphi\left(\frac{M-\mu}{\sigma_o}\right)\cdot\tilde{\mu}\cdot\tilde{\sigma}$$

$$\int_{-\infty}^{\infty} x^2\varphi\left(\frac{x-\mu}{s}\right)\varphi\left(\frac{M-x}{r}\right)dx = \varphi\left(\frac{M-\mu}{\sigma_o}\right)\cdot\tilde{\sigma}\cdot(\tilde{\mu}^2+\tilde{\sigma}^2)$$

Integration by parts gives

$$\int_{-\infty}^{\infty} x(x-\mu)\ \varphi\left(\frac{x-\mu}{s}\right)\Phi\left(\frac{M-x}{r}\right)dx = s^2\int_{-\infty}^{\infty}\varphi\left(\frac{x-\mu}{s}\right)\Phi\left(\frac{M-x}{r}\right)dx$$

$$-\frac{s^2}{r}\int_{-\infty}^{\infty} x\varphi\left(\frac{x-\mu}{s}\right)\varphi\left(\frac{M-x}{r}\right)dx$$

and thus using the above formulae

$$\int_{-\infty}^{\infty} x^2\varphi\left(\frac{x-\mu}{s}\right)\Phi\left(\frac{M-x}{r}\right)dx = s(\mu^2+s^2)\Phi\left(\frac{M-\mu}{\sigma_o}\right)-\left(\frac{s}{\sigma_o}\right)^3(2r^2\mu+(M+\mu)s^2)\varphi\left(\frac{M-\mu}{\sigma_o}\right)$$

ii) actual calculations

$$\mu_o^M(\theta) = \int_{-\infty}^{M}\frac{x}{r}\varphi\left(\frac{x-\theta}{r}\right)dx + M\cdot \Pr[X \ge M_{/\theta}]$$

$$= -r\varphi\left(\frac{M-\theta}{r}\right)+\theta\Pr[X\le M_{/\theta}] + M\Pr[X\ge M_{/\theta}]$$

$$= M + (\theta-M)\Phi\left(\frac{M-\theta}{r}\right) - r\varphi\left(\frac{M-\theta}{r}\right)$$

$$\mu_o(\theta) = \theta$$

Applying the formulae derived in i) we get by straightforward calculations

$$E[\mu_o^M(\theta)] = \int_{-\infty}^{\infty}\mu_o^M(\theta)\ \frac{1}{s}\varphi\left(\frac{\theta-m_o}{s}\right)d\theta$$

$$= M-(M-m_o)\Phi\left(\frac{M-m_o}{\sigma_o}\right)-\sigma_o\varphi\left(\frac{M-m_o}{\sigma_o}\right)$$

$$E[\theta\cdot\mu_o^M(\theta)] = Mm_o + (s^2+m_o^2-Mm_o)\Phi\left(\frac{M-m_o}{\sigma_o}\right)-m_o\sigma_o\varphi\left(\frac{M-m_o}{\sigma_o}\right)$$

Hence

$$\mathrm{Cov}[\mu_o^M(\theta),\mu_o(\theta)] = E[\theta\cdot\mu_o^M(\theta)] - m_o\, E[\mu_o^M(\theta)] = s^2\Phi\left(\frac{M-m_o}{\sigma_o}\right)$$

and

$$\boxed{b_1 = (1-\pi)^2 \, w\Phi\left(\frac{M-m_o}{\sigma_o}\right)} .$$

As $\mathrm{Cov}[X_1 \wedge M, \; X_2 \wedge M / X_1, X_2 \text{ ordinary}]$

$$= \mathrm{Cov}[U_1 \wedge M, \; U_2 \wedge M]$$

$$= E[\mathrm{Cov}[U_1 \wedge M, \; U_2 \wedge M/_\theta]] + \mathrm{Cov}[E[U_1 \wedge M/_\theta], \; E[U_2 \wedge M/_\theta]]$$

$= \mathrm{Var}[\mu_o^M(\theta)]$, we conclude from 19)

$$\boxed{b_2 = (1-\pi)^2 \, \mathrm{Cov}[U_1 \wedge M, \; U_2 \wedge M]} .$$

To obtain a closed formula for b_3, observe

$$\int_{-\infty}^{M} x(x-\mu) \frac{1}{\sigma} \varphi(x-\mu)\,dx = -x\sigma\varphi\left(\frac{x-\mu}{\sigma}\right)\Big|_{-\infty}^{M} + \sigma \int_{-\infty}^{M} \varphi\left(\frac{x-m}{\sigma}\right) dx$$

$$= -\sigma M \varphi\left(\frac{M-\mu}{\sigma}\right) + \sigma^2 \Phi\left(\frac{M-\mu}{\sigma}\right)$$

$$\int_{-\infty}^{M} x^2 \frac{1}{\sigma} \varphi(x-\mu)\,dx = -\sigma(M+\mu)\;\varphi\left(\frac{M-\mu}{\sigma}\right) + (\mu^2+\sigma^2)\Phi\left(\frac{M-\mu}{\sigma}\right)$$

According to 1) the density function of X is

$$f(x) = \int (1-\pi) p_o(x/_\theta)\; dU(\theta) + \pi p_e(x) = (1-\pi) p_o(x) + \pi p_e(x)$$

with (see 6.1 and 6.2)

$$p_o(x) = \frac{1}{\sigma_o} \varphi\left(\frac{x-m_o}{\sigma_o}\right)$$

$$p_e(x) = \frac{1}{\sigma_e} \varphi\left(\frac{x-\mu_e}{\sigma_e}\right)$$

Hence

$$\boxed{\begin{aligned} E[(X\wedge M)^2] &= (1-\pi)\left\{(m_o^2+\sigma_o^2)\,\Phi\left(\frac{M-m_o}{\sigma_o}\right) - \sigma_o(M+m_o)\,\varphi\left(\frac{M-m_o}{\sigma_o}\right)\right\} \\ &\quad + \pi\left\{(\mu_e^2+\sigma_e^2)\,\Phi\left(\frac{M-\mu_e}{\sigma_e}\right) - \sigma_e(M+\mu_e)\,\varphi\left(\frac{M-\mu_e}{\sigma_e}\right)\right\} \\ &\quad + M^2\left\{1-(1-\pi)\,\Phi\left(\frac{M-m_o}{\sigma_o}\right) - \pi\Phi\left(\frac{M-\mu_e}{\sigma_e}\right)\right\} \\ &= A \end{aligned}}$$

The same calculations as at the beginning of ii) leading to the formula for $\mu_o^M(\theta)$ are repeated to obtain $E[X\wedge M]$, of course with different parameter values. From this calculation we obtain

$$\boxed{\begin{aligned} E[X\wedge M] &= M + (1-\pi)\left\{(m_o-M)\,\Phi\left(\frac{M-m_o}{\sigma_o}\right) - \sigma_o\varphi\left(\frac{M-m_o}{\sigma_o}\right)\right\} \\ &\quad + \pi\left\{(\mu_e-M)\,\Phi\left(\frac{M-\mu_e}{\sigma_e}\right) - \sigma_e\varphi\left(\frac{M-\mu_e}{\sigma_e}\right)\right\} \\ &= B \end{aligned}}$$

We can now, finally, write

$$\boxed{b_3 = \mathrm{Var}[X\wedge M] = A - B^2}$$

FILMED

4–8